Introduction to Experimental Design

BIM 283

Advanced Design of Experiments for Biomedical Engineers

Contact Information

- David Rocke
  - dmrocke@ucdavis.edu
  - http://dmrocke.ucdavis.edu
  - (530) 304-1019 (cell)
- Brittany Lemmon
  - blemmon@ucdavis.edu

Class Meetings:	Tuesday and Thursday 10:00am–11:20am 2202 GBSF
Lab:	Tuesday and Thursday 11:20am–11:50am 2202 GBSF
Office Hours:	Tuesday 1:00pm–2:00pm, 140B Med Sci 1C Or by appointment, in person or on Zoom.
Office:	140B Med Sci 1C Cell: 530-304-1019 e-mail: dmrocke@ucdavis.edu dmrocke.ucdavis.edu/Class/BIM283.2024.Winter/BIM283-Winter-2024.html
	Email list: bim283-w24@ucdavis.edu Canvas site: BIM 283 001 WQ 2024

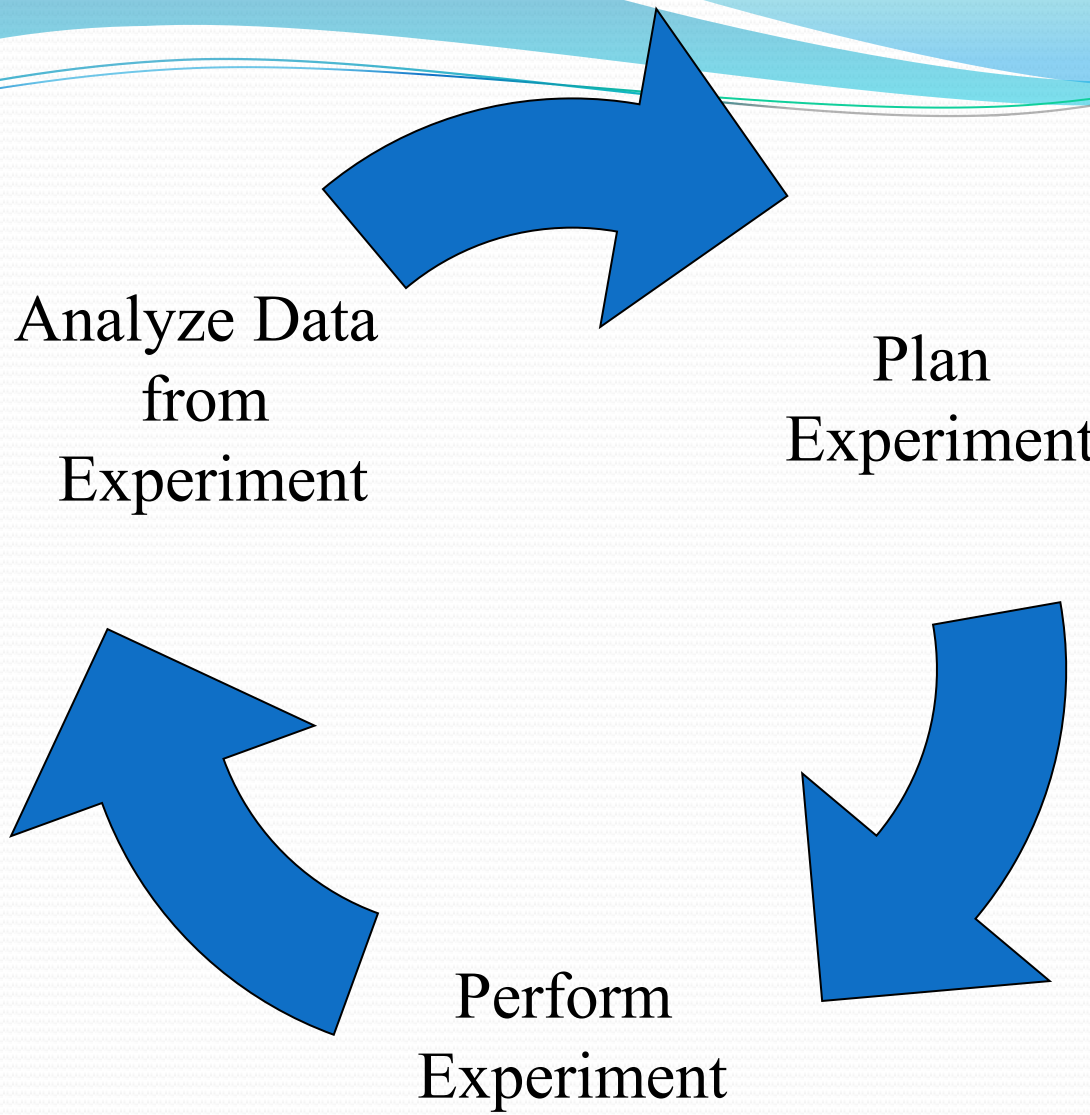
Required Text:	Statistics for Experimenters, Second Edition. Box, GEP, Hunter, JS, and Hunter, WG, Wiley, 2005.
Software:	Lectures and homework will utilize R for computation.
TA:	Brittany Lemmon (blemmon@ucdavis.edu).
Course Grading:	Letter Grades based on – Homework – Exams – Possible Projects – Students should attend class unless they are ill – Lecture slides will be posted but class will not be recorded
Prerequisites	It is assumed that the student has taken at least one introductory statistics class.

Basic Principles of Experimental Investigation

- Sequential Experimentation
- Comparison
- Manipulation
- Replication
- Randomization
- Blocking
- Simultaneous variation of factors
- Main effects and interactions
- Sources of variability

Sequential Experimentation

- No single experiment is definitive
- Each experimental result suggests other experiments
- Scientific investigation is iterative.
- "No experiment can do everything; every experiment should do something," *George Box*.


Analyze Data from Experiment
Plan Experiment
Perform Experiment

Comparison

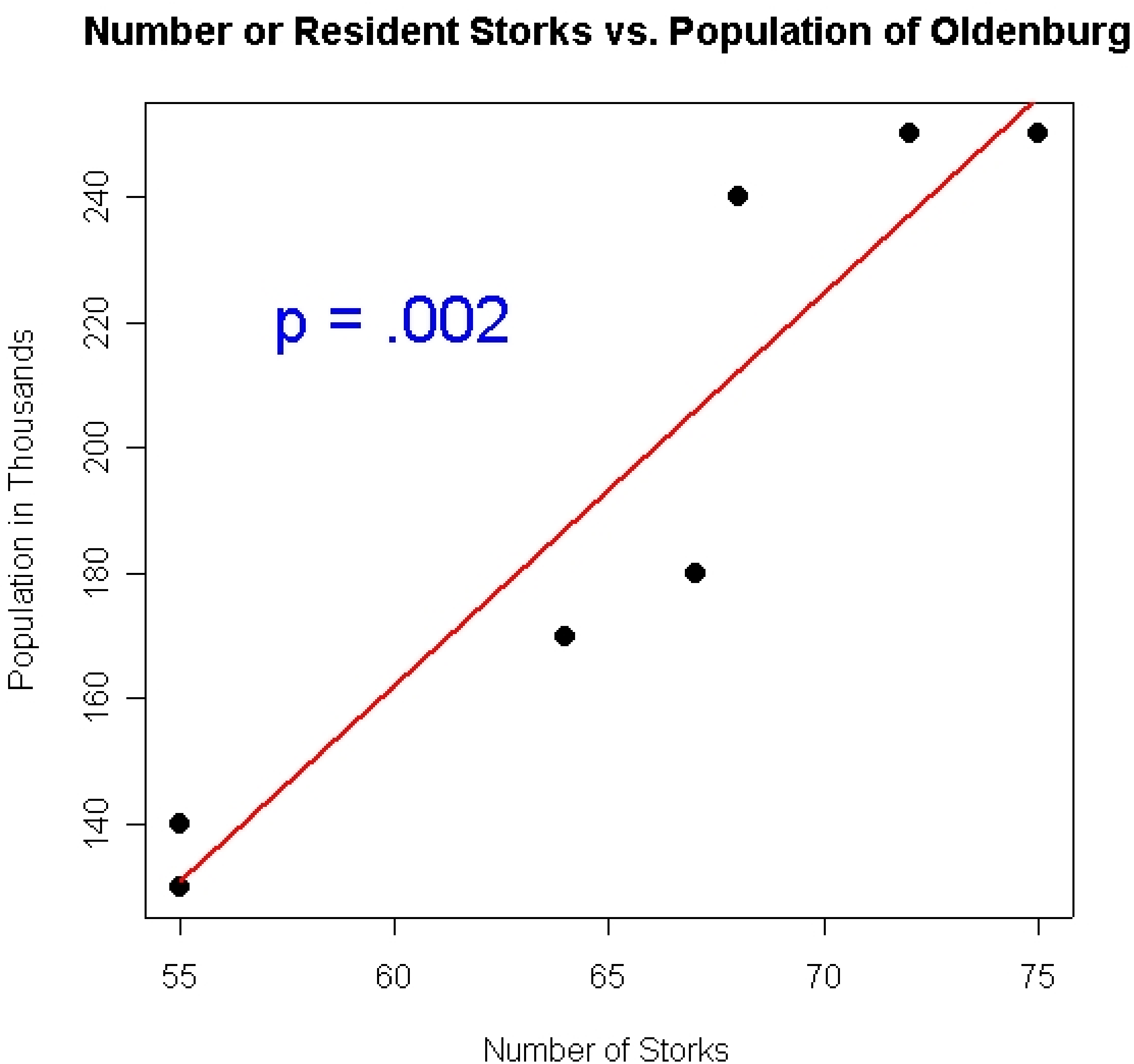
- Usually absolute data are meaningless, only comparative data are meaningful
- The measured level of mRNA for HNF-1 in a sample of liver cells is not meaningful
- The comparison of the measured mRNA levels of HNF-1 in samples from normal and diseased liver cells might be meaningful

Internal vs. External Comparison

- Comparison of an experimental results with historical results is likely to mislead
- Many factors that can influence results other than the intended treatment
- Best to include controls or other comparisons in each experiment
- This may not be possible in clinical or observational studies, but is crucial in any laboratory study

Manipulation

- Different experimental conditions need to be imposed by the experimenters, not just observed, if at all possible
- The rate of complications in cardiac artery bypass graft surgery may depend on many factors which are not controlled (for example, characteristics of the patient), and may be hard to measure


Number or Resident Storks vs. Population of Oldenburg
p = .002
Population in Thousands
140
160
180
200
220
240
55
60
65
70
75
Number of Storks

Randomization

- Randomization limits the difference between groups that are due to irrelevant factors
- Such differences will still exist, but can be quantified by analyzing the randomization
- This is a method of controlling for unknown confounding factors

- Suppose that 50% of a patient population is female
- A sample of 100 patients will not generally have exactly 50% females
- Numbers of females between 40 and 60 would not be surprising
- In two groups of 100, the disparity between the number of females in the two groups can be as big as 20% simply by chance, but not much larger
- This also holds for factors we don’t know about

- Randomization does not exactly balance against any specific factor
- To do that one should employ *blocking*
- Instead, it provides a way of quantifying possible imbalance even of unknown factors
- Randomization even provides an automatic method of analysis that depends on the design and randomization technique.

The Farmer from Whidbey Island

- Visited the University of Washington with a Whalebone water douser
- 10 Dixie cups, 5 with water, 5 empty, each covered with plywood
- Placed in a random order defined by generating 10 random numbers and sorting the cups by the random number
- If he gets all 10 right, is chance a reasonable explanation?

$$\binom{10}{5} = 252$$

$$\frac{1}{252} = .004$$

- The randomness is produced by the process of randomly choosing which 5 of the 10 are to contain water
- There are no other assumptions

- If the randomization had been to flip a coin for each of the 10 cups, then the probability of getting all 10 right by chance is different
- There are $2^{10} = 1024$ ways for the randomization to come out, only one of which is corresponds to the choices, so the chance is $1/1024 = .001$
- Compare this to the 252 cases where there are exactly 5 of each type
- The method of randomization matters
- If the farmer could observe condensation on the cups, then this is still evidence of non-randomness, but not evidence of the effectiveness of dousing!

Randomization Inference

- 20 tomato plants are divided 10 groups of 2 placed next to each other in the greenhouse (to control for temperature and insolation)
- In each group of 2, one is chosen using a random number table to receive fertilizer A; the other receives fertilizer B
- The yield of each plant in pounds of tomatoes is measured
- The null hypothesis is that the fertilizers are equal in promoting tomato growth

	1	2	3	4	5	6	7	8	9	10
A	132	82	109	143	107	66	95	108	88	133
B	140	88	112	142	118	64	98	113	93	136
diff	8	6	3	-1	11	-2	3	5	5	3

Pounds of yield of tomatoes for 20 plants

- The average yield for fertilizer A is 106.3 pounds
- The average yield for fertilizer B is 110.4 pounds
- The average difference is 4.1
- Could this have happened by chance?
- Is it statistically significant? (Defined as not likely to have happened by pure chance alone.)
- If A and B do not differ in their effects (null hypothesis is true), then the plants' yields would have been the same either whether A or B is applied
- Each difference would be the negative of what it currently is, if the coin flip had come out the other way

Actual

Fert A $\downarrow$ 132 lb

$\Delta = 8$

Fert B $\downarrow$ 140 lb

Hypothetical

Fert B $\downarrow$ 132 lb

$\Delta = -8$

Fert A $\downarrow$ 140 lb

- In pair 1, the yields were 132 and 140.
- The difference was 8, but it could have been –8
- With 10 coin flips, there are $2^{10} = 1024$ possible outcomes of + or – on the difference
- These outcomes are possible outcomes from our action of randomization, and carry no assumptions
- The measurements don't have to be normally distributed or have the same variance

- Of the 1024 possible outcomes that are all equally likely under the null hypothesis, only 3 had greater values of the average difference, and only four (including the one observed) had the same value of the average difference
- There are also an equal number of cases where the outcome is the same magnitude but opposite sign.
- The likelihood of an outcome this extreme happening by chance is $2[3+4/2]/1024 = .0098$
- This does not depend on any assumptions other than that the randomization was correctly done

	1	2	3	4	5	6	7	8	9	10
A	132	82	109	143	107	66	95	108	88	133
B	140	88	112	142	118	64	98	113	93	136
diff	8	6	3	-1	11	-2	3	5	5	3

Paired t-test

$$\overline{d} = 4.1$$

$$s_d = 3.872$$

$$t_9 = \frac{4.1}{3.872/\sqrt{10}} = \frac{4.1}{1.224} = 3.35$$

$p = .0085$ (two-sided) by t-test

$p = .0098$ by true randomization distribution

same range for simulation randomization distributions

The t-test can be thought of as an approximation to the randomization distribution.

Replication

- Both for randomization inference and for the t-test, it is important to have enough replicates (in this case pairs of plants).
- With 6 pairs of plants, it is impossible to get a p-value less than 0.05 by randomization, the minimum number of pairs is 7.
- With the t-test, none of the sets of 2 pairs of plants from the data has a p-value less than 0.05. For the two most extreme differences, 11 and 8, the p-value is 0.0997.
- There are 6 sets of 3 difference with a p-value less than 0.05, but only 1/20 random choices of 3 satisfy this.

With the same mean difference and standard deviation of the difference

$$\bar{d} = 4.1$$

$$s_d = 3.872$$

$$t_{n-1} = \frac{4.1}{3.872 / \sqrt{n}} = 1.059\sqrt{n}$$

n has to be at least 5 for the p-value to be less than 0.05

It's all about the denominator!

And n is a big part of the denominator!

	1	2	3	4	5	6	7	8	9	10
A	132	82	109	143	107	66	95	108	88	133
B	140	88	112	142	118	64	98	113	93	136
diff	8	6	3	-1	11	-2	3	5	5	3

Randomization in practice

- Whenever there is a choice, it should be made using a formal randomization procedure, such as Excel's rand() function or the randomization functions in R.
- This protects against unexpected sources of variability such as day, time of day, operator, reagent, etc.
- In the lab, this may seem like a lot of trouble, but it is less trouble than dealing with false or irreproducible results.

Pair Number	First Sample Treatment
1	A or B?
2	A or B?
3	A or B?
4	A or B?
5	A or B?
6	A or B?
7	A or B?
8	A or B?
9	A or B?
10	A or B?

Pair Num	First Sample Treatment	random number
1	A or B?	0.871413
2	A or B?	0.786036
3	A or B?	0.889785
4	A or B?	0.081120
5	A or B?	0.297614
6	A or B?	0.540483
7	A or B?	0.824491
8	A or B?	0.624133
9	A or B?	0.913187
10	A or B?	0.001599

- =rand() in first cell
- Copy down the column
- Highlight entire column
- ^c (Edit/Copy)
- Edit/Paste Special/Values
- This fixes the random numbers so they do not recompute each time
- =IF(C3<0.5,"A","B") goes in cell C2, then copy down the column

Plant Pair	First Plant Treatment	random number
1	B	0.871413
2	B	0.786036
3	B	0.889785
4	**A**	**0.081120**
5	**A**	**0.297614**
6	B	0.540483
7	B	0.824491
8	B	0.624133
9	B	0.913187
10	**A**	**0.001599**

- To randomize run order, insert a column of random numbers, then sort on that column
- More complex randomizations require more care, but this is quite important and worth the trouble
- Randomization can be done in Excel, R, or anything that can generate (pseudo) random numbers

Randomization in R

Randomly select A or B for each unit

```
> treatments <- c("Treatment.A","Treatment.B")
> treats <- sample(treatments,10,replace=T)
> treats
 [1] "Treatment.B" "Treatment.B" "Treatment.A" "Treatment.A" "Treatment.B"
 [6] "Treatment.A" "Treatment.B" "Treatment.A" "Treatment.A" "Treatment.A"
```

Randomly assign 5 each of A and B to the 10 units

```
> treatments2 <- rep(treatments,each=5)
> treatments2
 [1] "Treatment.A" "Treatment.A" "Treatment.A" "Treatment.A" "Treatment.A"
 [6] "Treatment.B" "Treatment.B" "Treatment.B" "Treatment.B" "Treatment.B"
> treats2 <- sample(treatments2)
> treats2
 [1] "Treatment.B" "Treatment.A" "Treatment.A" "Treatment.B" "Treatment.B"
 [6] "Treatment.B" "Treatment.A" "Treatment.B" "Treatment.A" "Treatment.A"
```

What is Random?

- Physical randomization, like shuffling cards, requires a lot of shuffling and care to avoid systematic effects.
- Alpha decay of uranium 235 to thorium 231 and an alpha particle occurs at a random time with half-life 704 million years per atom. This is as far as we know, truly random.
- We use pseudo-random numbers in computation.
- These are produced by a deterministic algorithm, starting from a pre-set or user-set seed.
- The function `sample()` uses this pseudo-random number generator.

```
> runif(4)
[1] 0.1455781 0.1422823 0.4182901 0.3103100
> runif(4)
[1] 0.004705293 0.652554638 0.593491534 0.056253182
> set.seed(2022)
> runif(4)
[1] 0.8159777 0.6472593 0.1203286 0.5438002
> set.seed(2022)
> runif(4)
[1] 0.8159777 0.6472593 0.1203286 0.5438002
```

It is good practice to use set.seed() in any code that uses pseudo-random numbers.

Randomization in R

Randomly select A or B for each unit

```
> treatments <- c("Treatment.A","Treatment.B")
> set.seed(2022)
> treats <- sample(treatments,10,replace=T)
> treats
 [1] "Treatment.B" "Treatment.A" "Treatment.B" "Treatment.A" "Treatment.A"
 [6] "Treatment.B" "Treatment.B" "Treatment.A" "Treatment.B" "Treatment.B"
```

Four "A" and six "B"

Randomly assign 5 each of A and B to the 10 units

```
> treatments2 <- rep(treatments,each=5)
> treatments2
 [1] "Treatment.A" "Treatment.A" "Treatment.A" "Treatment.A" "Treatment.A"
 [6] "Treatment.B" "Treatment.B" "Treatment.B" "Treatment.B" "Treatment.B"
> set.seed(2022)
> treats2 <- sample(treatments2)
> treats2
 [1] "Treatment.A" "Treatment.A" "Treatment.B" "Treatment.B" "Treatment.B"
 [6] "Treatment.B" "Treatment.A" "Treatment.B" "Treatment.A" "Treatment.A"
```